

Claims

1. A data processing method for extracting a subset as a processing object from a tabular format data expressed as an array of records each including an item and an item value belonging to the item, comprising:

a) a step of constructing the tabular format data by dividing it into information blocks each including a value list in which item values are stored in order of item value numbers corresponding to the item values belonging to a specific item and a pointer array in which pointer values to indicate the item value numbers are stored in order of unique record numbers;

b) a step of creating, in response to selection of a subset as a part of the records, an ordered set array containing recording numbers of the subset;

c) a step of arranging a pointer value in the pointer array at a position indicated by each of the record numbers of the ordered set into an item value number array correspondingly to the position where the record number is arranged; and

d) a step of creating a second pointer array storing position elements and a second value list storing value elements by referring to the value in the item value number array, and wherein

a value in the value list is specified from a record number

of the ordered set array through an element in the second pointer array at a position indicted by the record number and an element in the second value list at a position indicated by the element in the second pointer array.

2. The data processing method according to claim 1, characterized in that the step of creating the second pointer array and the second value list includes:

a step of sorting the elements in the item value number array, creating an array in a state where a duplicate value is excluded, and making this the second value list; and

a step of creating the second pointer array by converting the elements in the item value number array to reflect the sort and the exclusion of the duplicate value.

3. A data processing method for extracting a subset as a processing object from a tabular format data expressed as an array of records each including an item and an item value belonging to the item, comprising:

a) a step of constructing the tabular format data by dividing it into information blocks each including a value list in which item values are stored in order of item value numbers corresponding to the item values belonging to a specific item and a pointer array in which pointer values to indicate the item value numbers are stored in order of unique record numbers;

b) a step of creating, in response to selection of a subset as a part of the records, an ordered set array containing

recording numbers of the subset;

c) a step of arranging a pointer value in the pointer array at a position indicated by each of the record numbers of the ordered set into an item value number array correspondingly to the position where the record number is arranged; and

e) a step of using the item value number array as a second value list and using the original value list as a second value list, and wherein

a value in the value list is specified from a record number of the ordered set array through an element in the second pointer array at a position indicated by the record number and an element in the second value list at a position indicated by the element in the second pointer array.

4. The data processing method according to any one of claims 1 to 3, characterized by further comprising a step of comparing a size of the ordered set array with a size of the value list,

wherein as a result of the comparison, in a case where the size of the subset is smaller than the size of the value list at a predetermined rate, the steps c) and d) or the steps c) and e) are performed.

5. A retrieval method using the second pointer array and/or the second value list created by the data processing method according to any one of claims 1 to 4, the method

characterized by comprising:

a step of specifying an element indicating an item value as a retrieval object among elements in the second value list;

a step of arraying a value indicating that a flag is on at a position corresponding to a position of the specified element in a flag array having the same size as the second value list;

a step of specifying an element of the second pointer array indicated by a record number in the ordered set;

a step of referring to a state of a flag at a position indicated by an element of the second pointer array in the flag array; and

a step of successively arranging the record number into a newly provided ordered set array for output in a case where the state of the flag is on.

6. An aggregation method using the second pointer array and/or the second value list created by the data processing method according to any one of claims 1 to 4, the method characterized by comprising:

a step of creating a classification number array in which a classification number indicating a category of a value is arranged correspondingly to an element of the second value list;

a step of specifying a record number in the ordered set array corresponding to the specified classification number;

and

a step of performing aggregation using a predetermined value list item value indicated by the specified record number.

7. A sort method using the second pointer array and/or the second value list created by the data processing method according to any one of claims 1 to 4, the method characterized by comprising:

a step of calculating an existence number as the number of elements indicated by the second pointer array for each value of the second value list;

a step of creating, based on the existence number, a cumulative number array corresponding to a value of the second value list and indicating a head position at which a record number in the ordered set array is to be arranged; and

a step of referring to the cumulative array and arranging a record number of the ordered set array into an array for output so that a sort order of the item values in the value list is reflected.

8. A method of joining tabular format data by using the second pointer array and/or the second value list in the information block relating to each of plural tabular format data created by the data processing method according to any one of claims 1 to 4, the method characterized by comprising:

a step of finding an item to be shared in each of the plural tabular format data;

a step of equating item values in the second value list of the information block relating to the item; and

a step of, in response to equating the item values, updating an element in the second pointer array in each of the information blocks in accordance with a change in arrangement of the item values.

9. A data processing program for operating a computer to extract a subset as a processing object from a tabular format data expressed as an array of records each including an item and an item value belonging to the item, the data processing program causing the computer to execute:

a) a step of constructing the tabular format data by dividing it into information blocks each including a value list in which item values are stored in order of item value numbers corresponding to the item values belonging to a specific item and a pointer array in which pointer values to indicate the item value numbers are stored in order of unique record numbers;

b) a step of creating, in response to selection of a subset as a part of the records, an ordered set array containing recording numbers of the subset;

c) a step of arranging a pointer value in the pointer array at a position indicated by each of the record numbers of the ordered set into an item value number array correspondingly to the position where the record number is arranged; and

d) a step of creating a second pointer array storing position elements and a second value list storing value elements by referring to the value in the item value number array, and wherein

the computer is operated such that a value in the value list is specified from a record number of the ordered set array through an element in the second pointer array at a position indicted by the record number and an element in the second value list at a position indicated by the element in the second pointer array.

10. The data processing program according to claim 9, characterized in that

in the step of creating the second pointer array and the second value list, the computer is made to execute:

a step of sorting the elements in the item value number array, creating an array in a state where a duplicate value is excluded, and making this the second value list; and

a step of creating the second pointer array by converting the elements in the item value number array to reflect the sort and the exclusion of the duplicate value.

11. A data processing program for operating a computer to extract a subset as a processing object from a tabular format data expressed as an array of records each including an item and an item value belonging to the item, the data processing program causing the computer to execute:

a) a step of constructing the tabular format data by dividing it into information blocks each including a value list in which item values are stored in order of item value numbers corresponding to the item values belonging to a specific item and a pointer array in which pointer values to indicate the item value numbers are stored in order of unique record numbers;

b) a step of creating, in response to selection of a subset as a part of the records, an ordered set array containing recording numbers of the subset;

c) a step of arranging a pointer value in the pointer array at a position indicated by each of the record numbers of the ordered set into an item value number array correspondingly to the position where the record number is arranged; and

e) a step of using the item value number array as a second value list and using the original value list as a second value list, and wherein

the computer is made to operated such that a value in the value list is specified from a record number of the ordered set array through an element in the second pointer array at a position indicted by the record number and an element in the second value list at a position indicated by the element in the second pointer array.

12. The data processing program according to any one of claims 9 to 11, characterized in that the computer is made

to execute a step of comparing a size of the ordered set array with a size of the value list, and

as a result of the comparison, in a case where the size of the subset is smaller than the size of the value list at a predetermined rate, the computer is made to execute the steps c) and d) or the steps c) and e).